KISII UNIVERSITY

UNIVERSITY EXAMINATIONS

FEB 2022-JUNE 2022

TIME: 3 hours

DPBS 931– Advanced Biostatistics

INSTRUCTIONS

1. Do not write anything on this paper
2. Answer ALL questions.
3. Show ALL your workings and/or steps of analysis.

1. (a)Explain the following Terminologies [2 marks]

   (i)Data

   (ii) Random Variable

   (iii)Sample

   (iv)Population.

   (b)Discuss the following [2 marks]

   (i) sensitivity

   (ii)predictive value negative

   (iii) Conditional probability

   (iv) Bayes' theorem

   (c)The following data shows the body mass index (BMI) for craniofacial morphology of patients diagnosed with obstructive sleep apnoea syndrome (OSAS).

38.57	27.78	40.81
38.34	29.01	47.78
26.86	54.33	28.99
25.21	30.49	27.38
36.42	41.50	29.39
24.54	41.75	44.68
24.49	33.23	47.09
29.07	28.21	42.10
26.54	27.74	33.48
51.44	30.08	

   Use these data to construct a frequency distribution and histogram. [6 marks]

2. The following data shows the gestational age of 50 fetuses that underwent open fetal myelomeningocele closure .

25	25	26	27	29	29	29	30	30	31
32	32	32	33	33	33	33	34	34	34
35	35	35	35	35	35	35	35	35	36
36	36	36	36	36	36	36	36	36	36
36	36	36	36	36	36	36	36	37	37

(i) construct a stem and leaf plot for these gestational ages [3 marks]

(ii) Based on the stem and leaf plot, what one word would you use to describe the nature of the data? [1 mark]

(iii) Why do you think the stem and leaf plot looks the way it does? [1 mark]

(iv) Compute the mean, median, mode, variance and standard deviation [5 marks].

3. The following shows findings in a study between self reported cancer cases and actual cases of breast cancer of a given country in Africa.

	Cancer cases		
Cancer reported (A)	Cancer in registry (B)	Cancer not in registry	Total
Yes	2991	2244	5235
No	112	115849	115961
Total	3103	118093	121196

(a)Let A be the event of reporting breast cancer. Find the probability of A in this study [1 mark]

(b)Let B be the event of having breast cancer confirmed in registry. Find the probability of B in this study [1 mark]

(c) Find $P(B|A)$ [2 marks]

(d) Find the sensitivity using self reported breast cancer as a predictor of actual breast cancer in the registry. [3 marks]

(e) Find the specificity of using self reported cancer as a predictor of actual breast cancer in the registry[3 marks].

4. (a) In a study between measles vaccination and Giuliani-Barres syndrome (GBS), a Poisson model was used in the examination of GBS during latent periods after vaccination. It was found that during the latent period, the rate of GBS was 1.28 cases per day. Using this estimate rounded to 1.3, find the probability on a given day of:

(i) No cases of GBS [1 marks]

(ii) At least one case of GBS [2 marks]

(iii) Fewer than five cases of GBS [2 marks]

(b) A nurse supervisor found that staff nurses, on average, complete a certain task in 10 minutes.If the times required to complete a task are approximately normally distributed with a standard deviation of 3 minutes, find:

(i) The proportion of nurses completing the task in less than 4 minutes [1 marks]

(ii) The proportion of nurses requiring more than 5 minutes to complete the task[2 marks]

(iii) The probability that a nurse who has just been assigned the task will complete it within 3 minutes [2 marks].

5. (a) (i)How does the sampling distribution of the sample mean, when sampling is with replacement, differ from sampling obtained when sampling is with replacement? [1 mark]

(ii) Describe the sampling distribution of the sample proportion when large samples are drawn [1 mark]

(b) (i) what are the assumptions underlying the use of the t distribution in estimating the difference between two population means? '[2 marks]

(ii) Arterial blood gas analyses performed on a sample of 15 physically active adult males yielded the following resting PaO_2 values:

75,80,80,74,84,78,89,72,83,76,75,87,78,79,88.

Compute a 95% confidence interval for the mean of the population.[2 marks]

(c) Explain the following: [2 marks]

(i) Type I error

(ii) Type II error

(d) Circulating levels of estrone were measured in a sample of 25 postmenopausal women following estrogen treatment. The sample mean and standard deviation were 73 and 16, respectively. At the 0.05 significance level can one conclude on the basis of these data that the population mean is higher than 70? [2 marks]

6. The following are the pulmonary blood flow (PBF) measured in ml/sqM , denoted by X, and Pulmonary blood volume (PBV) values recorded for 16 infants and children with congenial heart disease.

Y	X
168	4.31
280	3.40
420	17.80
303	12.30
429	13.99
605	8.73
522	8.90
224	5.87
291	5.00
233	3.51
370	4.24
531	19.41
516	16.61
211	7.21
439	11.60

find the regression equation describing the linear relationship between the two variables and test $H_0 : \beta = 0$ using the t test. Let $\alpha = 0.05$ [10 marks]